Application of the EEPAS Model to Forecasting Earthquakes of Moderate Magnitude in Southern California

David A. Rhoades
GNS Science

INTRODUCTION

The EEPAS (Every Earthquake a Precursor According to Scale) model is a method of long-range forecasting that uses the previous minor earthquakes in a catalog to forecast the major ones. It is based on the precursory scale increase (Ψ) phenomenon, which involves an increase in the magnitude and rate of occurrence of minor earthquakes close to the source region of a major event in preparation, including most recent major earthquakes in California (Evison and Rhoades 2002, 2004). The period of time occupied by the increase scales with magnitude, but it is on the order of 15 years for an **M** 7 event, five years for an **M** 6 event, and one or two years for an **M** 5 event. With a one-year time horizon as specified for the Regional Earthquake Likelihood Models (RELM) testing in southern California, it is therefore feasible to consider using the model to forecast earthquakes of **M** 5 and above.

The model has previously been fitted to the New Zealand earthquake catalog using earthquakes of magnitudes exceeding 3.95 to forecast those exceeding magnitude 5.75. It was shown to explain the data much better than a baseline model that is in principle time-invariant and has a location distribution based on proximity to the epicenters of past earthquakes. In the same form, and with the same magnitude thresholds, it was tested on California over the period 1975–2001 and again performed much better than the baseline model (Rhoades and Evison 2004). In the same form, but with magnitude thresholds one unit higher, it was tested on Japan over the period 1965–2001 and produced a similar result (Rhoades and Evison 2005), albeit with a smaller advantage over the baseline model. In order to fit well to lower magnitudes down to **M** 6.25, some adjustment of parameters, and in particular the magnitude-scaling parameter, was found to be necessary.

Recently, the model has been fitted at moderate magnitudes down to **M** 4.75 to the high-quality catalog of the Kanto region, central Japan, prepared by the National Research Institute for Earth Science and Disaster Prevention (NIED). In the Kanto region, earthquakes occur over a wide range of depths. The model was found to be informative down to 120 km. The results support the conclusion that the Ψ phenomenon is a feature of the preparation process not only of the largest earthquakes but also of most shallow earthquakes of moderate magnitude (Rhoades and Evison 2006). In the less-complex region of southern California, nearly all of the earthquakes are shallow, with more than 99% originating at depths less than 20 km. The catalog of southern California, like that of the Kanto region, is of high enough quality to support the fitting of the EEPAS model at moderate magnitudes.

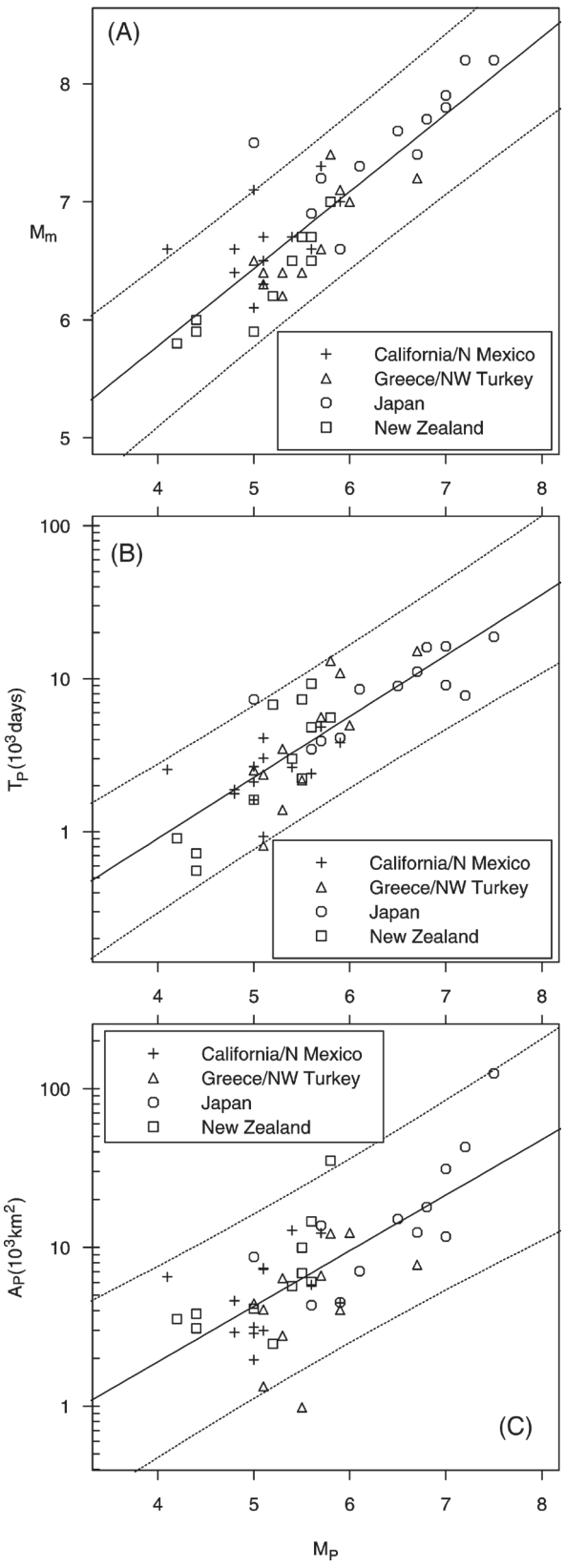
EEPAS MODEL

The EEPAS model adopts predictive scaling relations derived empirically from 47 examples of the Ψ phenomenon by Evison and Rhoades (2004) in the earthquake catalogs of California and northern Mexico, Japan, Greece, and New Zealand. These relations, shown in figure 1, are linear regressions of mainshock magnitude M_m, the logarithm of precursor time T_p, and the logarithm of precursor area A_p, on precursor magnitude M_p. There M_p is the average of the three largest magnitudes in the precursory period, T_p is the time between the onset of Ψ and the major earthquake, and A_p is the area occupied by the precursory earthquakes, major earthquake, and aftershocks. In the EEPAS model, the question of recognizing the Ψ phenomenon in advance of the major earthquake is set aside and every earthquake is regarded as a (possible) precursor of larger earthquakes to follow it in the long term. An earthquake's contribution to the earthquake-rate density depends on its magnitude, which is treated as an individual instance of M_p.

The rate density $\lambda(t,m,x,y)$ of earthquake occurrence is defined for any time t, magnitude m, and location (x,y), where m exceeds a threshold magnitude m_c and (x,y) is a point in a region of surveillance R. Each earthquake (t_i,m_i,x_i,y_i) contributes a transient increment $\lambda_i(t,m,x,y)$ to the future rate density in its vicinity, given by

$$\lambda_i(t,m,x,y) = w_i f_{1i}(t) g_{1i}(m) h_{1i}(x,y) \qquad (1)$$

where w_i is a weighting factor that may depend on other earthquakes in the vicinity, and f_{1i}, g_{1i} and h_{1i} are densities of the probability distributions for time, magnitude, and location, respectively. The magnitude density g_{1i} is assumed to take the form



▲ **Figure 1.** Predictive relations, derived from 47 examples of the Ψ phenomenon by Evison and Rhoades (2004). (A) Mainshock magnitude M_m on precursor magnitude M_P; (B) Precursor time T_P on M_P; (C) Precursor area A_P on M_P.

$$g_{1i}(m) = \frac{1}{\sigma_M \sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{m - a_M - b_M m_i}{\sigma_M} \right)^2 \right] \quad (2)$$

where a_M, $b_{M,}$ and σ_M are parameters (figure 1A). The time density f_{1i} is assumed to take the form

$$f_{1i}(t) = \frac{H(t - t_i)}{(t - t_i)\sigma_T \ln(10)\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{\log(t - t_i) - a_T - b_T m_i}{\sigma_T} \right)^2 \right] \quad (3)$$

where $H(s) = 1$ if $s > 0$ and 0 otherwise, and a_T, $b_{T,}$ and σ_T are parameters (figure 1B). The location density h_{1i} is assumed to take the form

$$h_{1i}(x, y) = \frac{1}{2\pi\sigma_A^2 10^{b_A m_i}} \exp\left[-\frac{(x - x_i)^2 + (y - y_i)^2}{2\sigma_A^2 10^{b_A m_i}} \right] \quad (4)$$

where σ_A and b_A are parameters (figure 1C). The total rate density is obtained by summing over all past occurrences, including earthquakes outside R, which could affect the rate density within R:

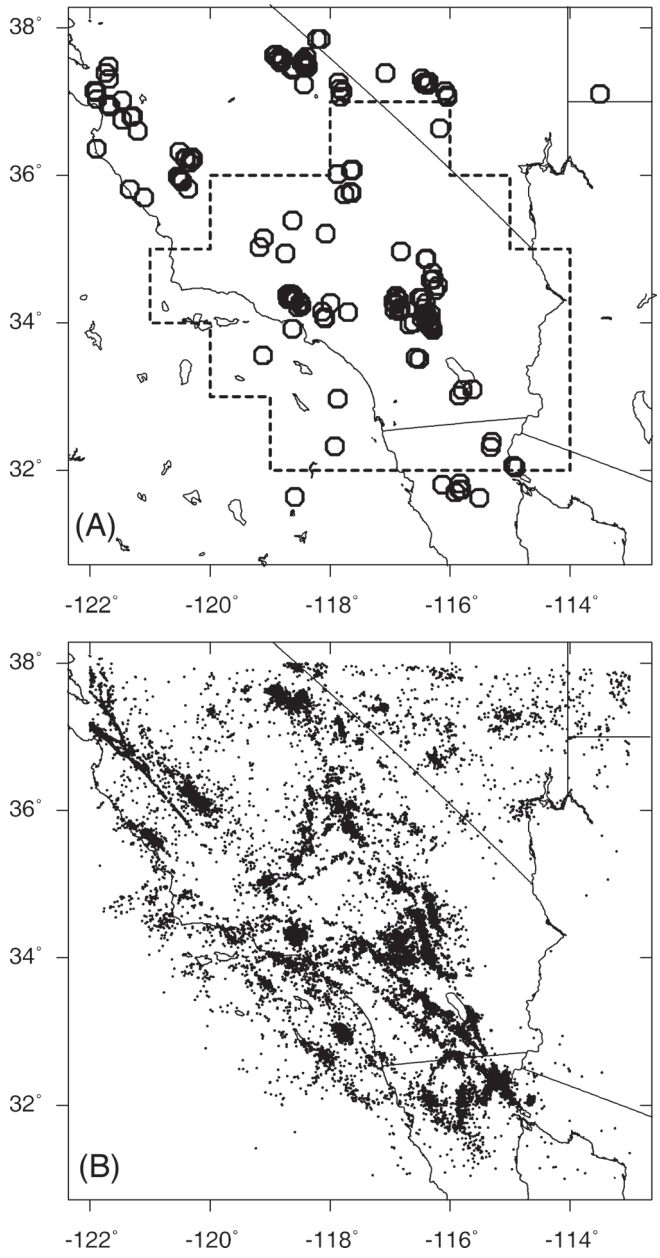
$$\lambda(t, m, x, y) = \mu\lambda_0(t, m, x, y) + \sum_{t_i \geq t_0; m_i \geq m_0} \eta(m_i)\lambda_i(t, m, x, y) \quad (5)$$

where μ is a parameter, λ_0 is a baseline rate density, t_0 is the time of the beginning of the catalog, and η is a normalizing function.

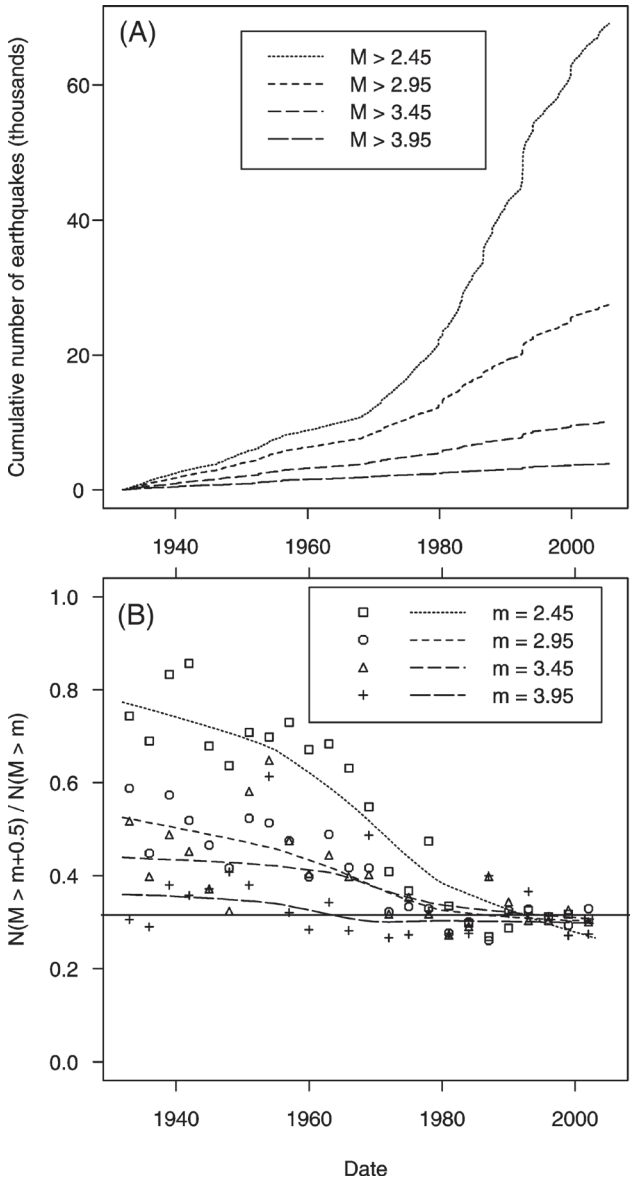
The overall form of the rate density (equation 5) is similar to that for epidemic-type models (Ogata 1989, 1998; Console and Murru 2001; Console *et al.* 2003; Jackson and Kagan 2000). Like the epidemic-type models, the structure is formally that of a branching process. However, conceptually there is a big difference; the idea here is not that a small earthquake triggers a larger one but that it provides evidence that a larger earthquake may be in preparation (Evison and Rhoades 2001). Also, the component distributions here do not follow power laws in time, location, or earthquake moment. The normal, lognormal, and bivariate normal distributions adopted in equations (2), (3), and (4), respectively, are chosen to be consistent with normally distributed errors in the predictive relations (figure 1). In the case of magnitudes, the standard frequency-magnitude relation is preserved overall by the normalizing function η, which makes the long-run average rate density under the EEPAS model, as a function of magnitude, *i.e.*, averaged over time and location, conform as well to the Gutenberg-Richter frequency-magnitude relation (Gutenberg and Richter 1944) as the catalog itself does.

When a lower magnitude threshold m_0 is applied for precursors, an adjustment must be made to compensate for the missing contribution from earthquakes with magnitudes below m_0. See Rhoades and Evison (2004) for details.

As in previous applications of the EEPAS model, the baseline model adopted here is fashioned from a model proposed by Jackson and Kagan (1999) in which the rate density λ_0 depends on proximity to past earthquakes (PPE) in the catalog with $\mathbf{M} > m_c$. For details of the PPE model, see Rhoades and Evison (2004, 2005).



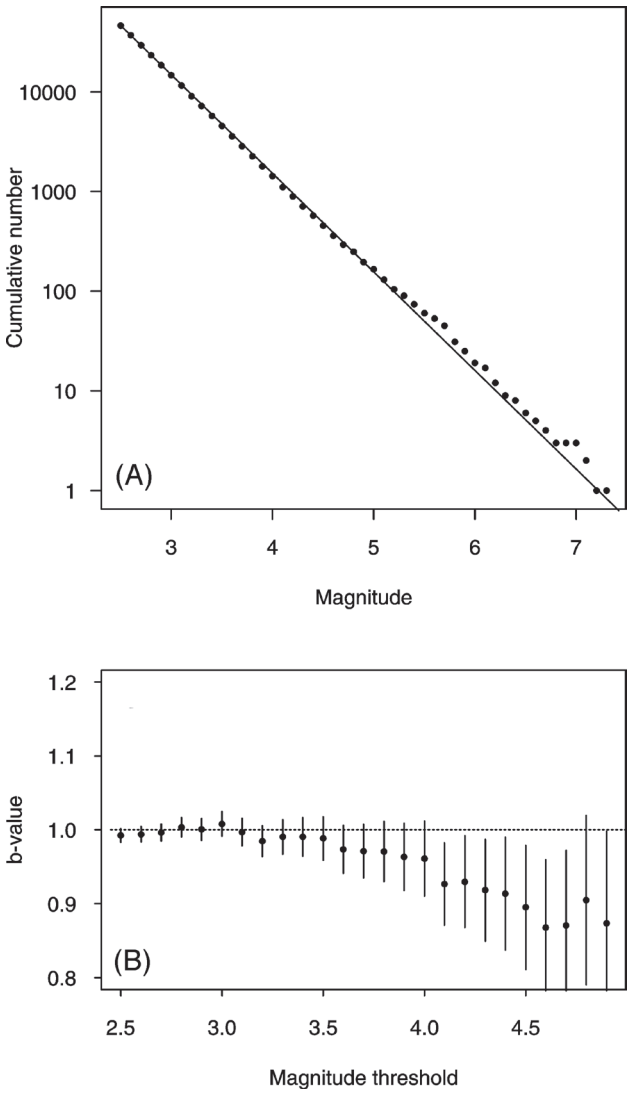
▲ **Figure 2.** Map of southern California showing (A) region of surveillance *R* (enclosed by dashed polygon) for the EEPAS model and epicenters of earthquakes with **M** > 4.95 during the period 1981–2004; (B) epicenters of earthquakes with **M** > 2.45 during the period 1981–2004.



▲ **Figure 3.** (A) Cumulative numbers of earthquakes within the whole region of figure 2 exceeding certain magnitude levels as a function of time; (B) Ratio of numbers of earthquakes $N(M > m + 0.5)$ / $N(M > m)$ in three-year intervals and smooth local regression fits. The solid line shows the expected value of the ratio under catalog completeness and a Gutenberg-Richter *b*-value of 1.

APPLICATION TO SOUTHERN CALIFORNIA

The EEPAS model was fitted to the Advanced National Seismic System (ANSS) Worldwide Earthquake Catalog, which is contributed to by members of the U.S. Council of the National Seismic System and available from the Northern California Earthquake Data Center, to maximize the likelihood of forecasting earthquakes over the period 1981–2004 within the region of surveillance *R* in southern California (figure 2A). To assess the suitability of the catalog, we examined the catalog of all earthquakes since the beginning of 1932 with **M** > 2.45 and with epicenters in the latitude range 31.0–38.0°N and longitude range 113.0–122.0°W. The distribution of epicenters is shown in figure 2(B). The cumulative number of earthquakes exceeding certain magnitude thresholds is shown in figure 3(A). The approximately constant slope of the curve for **M** > 2.45 since about 1980 indicates homogeneity of the catalog in this range. In order to judge the variation of magnitude completeness with time, ratios of the number of earthquakes exceeding certain magnitude thresholds have been plotted in figure 3(B). Let $N(M > m)$ be the number of earthquakes exceeding magnitude



▲ **Figure 4.** (A) Frequency magnitude relation for earthquakes in *R* in the learning period 1981–2004. The line shown has slope b = 0.99. (B) Maximum likelihood estimate of the Gutenberg-Richter b-value in *R* for 1981–2004, and its uncertainty (± 2 standard errors), as a function of minimum magnitude threshold.

m in a time interval. Under the assumption of catalog completeness and a Gutenberg-Richter b-value of 1, the expected value of the ratio $N(M > m + 0.5) / N(M > m)$ is 0.316, shown by the solid horizontal line in figure 3(B). If the catalog is adjudged to be sufficiently complete when this ratio drops below 0.4, then it is sufficiently complete for **M** > 3.95 from 1932, for **M** > 3.45 and **M** > 2.95 from about 1970, and for **M** > 2.45 from about 1980. On this basis, with a learning period of 1981–2004, the magnitude threshold for all earthquakes is set at $m_0 = 2.45$. This is adequately low, given that the typical difference in magnitude between precursory and major earthquakes is a little more than one unit (figure 1A) and that the magnitude threshold for targeted earthquakes is $m_c = 4.95$. There are 79 earthquakes exceeding m_c in the learning period, with epicenters inside *R*. The data from the earlier period 1932–1980 is needed for estimating parameters of the baseline model as well as providing data on precursory earthquakes.

The Gutenberg-Richter b-value is an important parameter for normalization of the EEPAS model. A frequency-magnitude plot of earthquakes within the region of surveillance over the period 1981–2004 is shown in figure 4(A), together with the maximum likelihood Gutenberg-Richter relation for **M** > 2.45. The graph is fairly linear for **M** < 5.0, with minor deviations above the line at higher magnitudes indicating a decrease in b-value. The maximum likelihood estimate of the b-value (Aki 1965) varies slightly with the choice of lower magnitude threshold (figure 4B) and takes values ranging from 1.01 to 0.87 with thresholds ranging from 2.45 to 4.85. For the EEPAS model, earthquakes above magnitude 4.0 contribute most to the rate density for magnitudes 5.0 and above. Hence the value $b = 0.96$, which is the estimate for a lower magnitude threshold of 4.0, is adopted.

The magnitude scaling parameter b_M was set to 1.0, which corresponds to perfect scaling of magnitudes and was also adopted for fitting of the model to the whole of Japan and the Kanto catalogs (Rhoades and Evison 2005, 2006). The parameters a_M, σ_M, a_T, b_T, σ_T, b_A, σ_A, and μ were then all fitted to maximize the likelihood of the catalog using the downhill simplex method (Nelder and Mead 1965). As in previous applications of the model, two weighting strategies were tried: the equal-weighting strategy $w_i = 1$ for all i, and a strategy to down-weight aftershocks. The latter strategy has proved superior in most previous applications, but here, as for the Kanto region with m_c = 4.75, the equal-weighting strategy is better. In this case it has a big advantage, which amounts to a gain of about 20 in the log likelihood. Therefore, only the results for the equal-weighting strategy are reported here.

As in most previous applications, the optimal value of μ is 0. With $b_M = 1$, $\mu = 0$, and $w_i = 1$, the normalization function reduces to the constant

$$\eta = \exp\left[-\beta\left(a_M + \frac{\sigma_M^2}{2}\right)\right] \qquad (6)$$

where $\beta = b\ln(10)$.

The optimal parameter values are listed in table 1. These values differ from those in previous applications in several respects. In particular, σ_M, b_T, b_A, and σ_T are relatively large, and a_M and σ_A are relatively small, compared with the corresponding values in previous applications. Thus the contribution that an earthquake of given magnitude makes to the future rate density has a magnitude distribution with a relatively low mean and large spread, a time distribution with a relatively high mean and large spread, and a location distribution with a relatively small spread.

A tentative explanation for these differences is offered, as follows. The magnitude threshold $m_c = 4.95$ is well below the largest earthquake in the catalog during the learning period (*i.e.*, the **M** 7.3 Landers mainshock of 28 June 1992). In such circumstances, a sizeable proportion of earthquakes exceeding m_c can occur as aftershocks of much larger events. For example, in the six months following the Landers mainshock there were 19

TABLE 1
EEPAS Model Parameter Values, Optimized for the Region of Surveillance, 1981–2004, Except as Noted.

Parameter	Value
m_0	2.45*
m_c	4.95*
b	0.96**
a_M	1.00
b_M	1.00*
σ_M	0.58
a_T	1.49
b_T	0.48
σ_T	0.81
b_M	0.61
σ_M	0.66
μ	0.00

* fixed
** see text

earthquakes of magnitude 5.0 and greater within 50 km of the epicenter. The EEPAS model was not designed to forecast aftershocks, but when it is formally optimized the parameter values will adjust to forecast both mainshock and aftershocks as well as possible. This explains the relatively low mean and large spread of the magnitude distribution. Now, the precursory earthquakes in the Ψ phenomenon usually constitute a Gutenberg-Richter set of magnitudes. The smaller magnitudes in this set would often be aftershocks of larger precursors, and if aftershocks were down-weighted their contribution would be minimized. But when many aftershocks have magnitudes exceeding m_c, the precursory aftershocks can be used to help forecast the aftershocks, if the time distribution for a given magnitude of precursor is lengthened. This explains the relatively high mean and large spread of the time distribution, as well the superiority of the equal-weighting strategy. The relatively small spread of the location distribution is attributed to the tight clustering of earthquake locations close to major faults in southern California (figure 2).

In view of the above considerations, the optimization of the parameters for $m_c = 4.95$ has probably compromised the ability of the model to forecast larger earthquakes with, say, **M** > 6, in order to improve the forecasting of the majority of the earthquakes in the target range, *i.e.*, those between **M** 5 and **M** 6. Nevertheless, the parameter values are the best to use for future RELM testing given the magnitude threshold, because the test period is also likely to include a similar proportion of aftershocks above the threshold.

There were $N = 79$ earthquakes with $M > m_c$ in the learning period. The optimal value of the log likelihood for the EEPAS model is −1339.5. The increase in log likelihood ($\Delta \ln L$) is 64.6 compared to the PPE baseline model and 160.8 compared to the model of least information, *i.e.*, a stationary and spatially uniform Poisson (SUP) model with Gutenberg-Richter magnitude distribution. This gives an information rate per earthquake ($\Delta \ln L / N$) of 0.82 compared to PPE and 2.04 compared to SUP. The former rate is higher than in the applications to New Zealand and Japan but similar to that for the whole of California with $m_c = 5.75$. The latter rate is higher than that for New Zealand, the whole of California, and the whole of Japan but slightly less than that for the Kanto region with $m_c = 4.75$.

CONCLUSION

The application of the EEPAS model to the catalog of southern California indicates that the model may be useful for long-range forecasting at moderate as well as large magnitudes in this region. The information rate per earthquake compared with time-invariant seismicity models is greater than or similar to that obtained in previous studies. The EEPAS model, as optimized for the period 1981–2004, will be submitted to the RELM testing exercise and used to make one-year forecasts, which will be updated annually.

This study also gives impetus to future research aimed at using the EEPAS model to further improve long- and short-term forecasts of moderate-to-large earthquakes in southern California, and perhaps to provide new models, or new variants of the present model, for RELM testing. First, there is the question of whether the model could be enhanced by adopting a modified magnitude distribution to explicitly allow for aftershocks of forecasted events. By this means, the present compromise in the parameter values of the time and magnitude distributions to accommodate aftershocks might be avoided. Secondly, although $\mu = 0$ in the present model, and hence no time-invariant baseline model is incorporated into the forecasts, it may be advantageous to EEPAS to incorporate some other time-invariant model more informative than PPE, if, as seems likely, such a model should become available during the RELM testing exercise. Finally, EEPAS could be used in lieu of a time-invariant baseline model for independent seismicity in conjunction with an epidemic-type short-term forecasting model to improve daily forecasts of earthquake probabilities. In such a combination, the EEPAS model would be expected to improve the forecasts of most mainshocks, except for those that have immediate foreshocks.

ACKNOWLEDGMENTS

This research was supported by the New Zealand Foundation for Research, Science and Technology under Contract No. C05X0402. Helpful reviews of the manuscript were provided by Russell Robinson, Warwick Smith, and Bob Simpson. The author is grateful to Matt Gerstenberger for providing information on the requirements for RELM testing.

REFERENCES

Aki, K. (1965). Maximum likelihood estimation of b in the formula log $N = a - bM$ and its confidence limits. *Bulletin of the Earthquake Research Institute, Tokyo University* **43**, 237–239.

Console, R., and M. Murru (2001). A simple and testable model for earthquake clustering. *Journal of Geophysical Research* **106**, 8,699–8,711.

Console, R., M. Murru, and A. M. Lombardi (2003). Refining earthquake clustering models. *Journal of Geophysical Research* **108**, 2,468, doi:10.1029/2002JB002130.

Evison, F., and D. Rhoades (2001). Model of long-term seismogenesis. *Annali di Geofisica* **44**, 81–93.

Evison, F., and D. Rhoades (2002). Precursory scale increase and long-term seismogenesis in California and northern Mexico. *Annals of Geophysics* **45**, 479–495.

Evison, F. F., and D. A. Rhoades (2004). Demarcation and scaling of long-term seismogenesis. *Pure and Applied Geophysics* **161**, 21–45.

Gutenberg, B., and C. F. Richter (1944). Frequency of earthquakes in California. *Bulletin of the Seismological Society of America* **34**, 185–188.

Jackson, D. D., and Y. Y. Kagan (1999). Testable earthquake forecasts for 1999. *Seismological Research Letters* **70**, 393–403.

Kagan, Y. Y., and D. D. Jackson (2000). Probabilistic forecasting of earthquakes. *Geophysical Journal International* **143**, 438–453.

Nelder, J. A., and R. Mead (1965). A simplex method for function minimization. *Computer Journal* **7**, 308–313.

Ogata, Y. (1989). Statistical models for standard seismicity and detection of anomalies by residual analysis. *Tectonophysics* **169**, 159–174.

Ogata, Y. (1998). Space-time point process models for earthquake occurrences. *Annals of the Institute of Statistical Mathematics* **50**, 379–402.

Rhoades, D. A., and F. F. Evison (2004). Long-range earthquake forecasting with every earthquake a precursor according to scale. *Pure and Applied Geophysics* **161**, 47–71.

Rhoades, D. A., and F. F. Evison (2005). Test of the EEPAS forecasting model on the Japan earthquake catalogue. *Pure and Applied Geophysics* **162**, 1,271–1,290.

Rhoades, D. A., and F. F. Evison (2006). The EEPAS forecasting model and the probability of moderate-to-large earthquakes in central Japan. *Tectonophysics* **417**, 119–130.

GNS Science
P. O. Box 30368
Lower Hutt
New Zealand
d.rhoades@gns.cri.nz